Form PTO-1390	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER P20294
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/700708

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/JP98/02302	26 May 1998	

TITLE OF INVENTION

METHOD FOR PREDICTING FUNCTIONS OF PROTEIN

APPLICANT(S) FOR DO/EO/US

Akiko ITAI, Nobuo TOMIOKA, Reiko ITAI

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information.

1. _X_ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.

2. ___ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.

3. _X_ This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)).

4. _X_ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).

5. _X_ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. _X_ is attached hereto (required only if not communicated by the International Bureau).
   b. _X_ has been communicated by the International Bureau.
   c. ___ is not required, as the application was filed in the United States Receiving Office (RO/US).

6. _X_ An English language translation of the International Application as filed (35 U.S.C. 371 (c)(2)).

7. _X_ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ___ are attached hereto (required only if not communicated by the International Bureau).
   b. ___ have been communicated by the International Bureau.
   c. ___ have not been made; however, the time limit for making such amendments has NOT expired.
   d. _X_ have not been made and will not be made.

8. ___ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3))

9. _X_ An oath or declaration of the inventor(s) (35 U.S.C. .371(c)(4)).
   "Unexecuted"

10. ___ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (U.S.C. 371(c)(5)).

Items 11 to 16 below concern other document(s) or information included:

11. Assignee: Institute of Medical Molecular Design, Inc.

12. ___ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.

13. ___ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.

14. _X_ A FIRST preliminary amendment.
    ___ A SECOND or SUBSEQUENT preliminary amendment.

15. ___ A substitute specification.

16. ___ A change of power of attorney and/or address letter.

17. _X_ Other items or information:
    Cover Sheet and International Application as published in Japanese.
    PCT/RO/101-PCT Request.
    PCT/IB/301.
    PCT/IB/308.
    PCT/IB/332.
    PCT/IPEA/408(in Japanese).
    PCT/IPEA/409(in Japanese).
    PCT/ISA/210(in English and Japanese).

526 Rec'd PCT/PTO 24 NOV 2000

U.S. APPLICATION NO. (If known, see 37 CFR 1.5) 09/700708	INTERNATIONAL APPLICATION NO. PCT/IPEA/02302	ATTORNEY'S DOCKET NUMBER P20294

18. ___ The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5)):		
Search report has been prepared by the EPO or JPO. . . . . $ 860.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482). . . . . $ 690.00		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO(37 CFR 1.445(a)(2). . . . . $ 710.00		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2) paid to USPTO. $1,000.00		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4). . . . . $ 100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT =	$860.00	
Surcharge of $130.00 for furnishing the oath or declaration later than ___ 20 ___ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$	

Claims	Number Filed	Number Extra	RATE		
Total Claims	8 - 20 =	0	X $18.00	$0.00	
Independent Claims	1 - 3 =	0	X $80.00	$0.00	
Multiple dependent claim(s) (if applicable)			+ $270.00	$0.00	
			TOTAL OF ABOVE CALCULATIONS =	$860.00	
Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by ½.				$	
			SUBTOTAL =	$860.00	
Processing fee of $130.00 for furnishing the English translation later than ___ 20 ___ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).			+		
			TOTAL NATIONAL FEE =	$860.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property			+		
			TOTAL FEES ENCLOSED =	$860.00	
			Amount to be refunded		$
			Charged		$

a. _X_ A check in the amount of $860.00 to cover the above fees is enclosed.

b. ___ Please charge my Deposit Account No. _____ in the amount of $_____ to cover the above fees.

c. _X_ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0089.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO **CUSTOMER NO. 7055**
AT THE PRESENT ADDRESS OF:
Bruce H. Bernstein
GREENBLUM & BERNSTEIN, P.L.C.
1941 Roland Clarke Place
Reston, VA 20191
(703) 716-1191

SIGNATURE
Bruce H. Bernstein, Reg No. 29,027
NAME
Reg No 33,094
29,027
REGISTRATION NUMBER

P20294.A02

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Akiko ITAI et al.

Examiner: Not known

Serial No : 09/700,708
(National Stage of PCT/JP98/02302)

I.A. Filed : May 26, 1998

Art Group: Not Known

For : METHOD FOR PREDICTING FUNCTIONS OF PROTEIN

PRELIMINARY AMENDMENT

Assistant Commissioner of Patents
Washington, D.C. 20231

Sir:

Prior to an examination of the above-identified patent application, the Examiner is respectfully requested to amend the application as follows:

IN THE SPECIFICATION

Please replace the Sequence Listing section filed with the application with the Sequence Listing being filed concurrently herewith.

REMARKS

The Examiner is respectfully requested to enter the foregoing amendment prior to examination of the above-identified patent application.

P20294.A02

If there are any comments or questions, the undersigned may be contacted at the below-listed telephone number.

Respectfully submitted,
Akiko ITAI et al.

Bruce H. Bernstein
Reg. No. 29,027

Reg No. 45905

August 21, 2001
GREENBLUM & BERNSTEIN, P.L.C.
1941 Roland Clarke Place
Reston, VA 20191
(703) 716-1191

P20294.P03

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant :A. ITAI et al.

Serial No : Not Yet Assigned

Filed : May 26, 1998

For :METHOD FOR PREDICTING FUNCTIONS OF PROTEIN

PRELIMINARY AMENDMENT

Commissioner of Patents and Trademarks
Washington, D.C. 20231

Sir:

Prior to the examination of the above-identified patent application, the Examiner is respectfully requested to amend the specification and claims as follows:

IN THE CLAIMS

Please amend the claims as follows:

In claim 3, line 1, delete "or claim 2".

In claim 4, line 1, change "any one of claims 1 through 3" to ---claim 1---.

In claim 5, line 2, change "any one of claims 1 through 3" to claim 1---.

In claim 7, line 1, delete "or claim 6".

REMARKS

The Examiner is respectfully requested to enter the foregoing amendment prior to examination and calculation of the filing fees in the above-identified patent application.

Should there be any questions, the Examiner is invited to contact the undersigned at the below listed number.

Respectfully submitted,
A. ITAI et al.

Bruce H. Bernstein
Reg. No. 29,027 Reg. No. 33,094

November 24, 2000
GREENBLUM & BERNSTEIN, P.L.C.
1941 Roland Clarke Place
Reston, VA 20191
(703) 716-1191

Method for Predicting Functions of Protein

Technical Field

This invention relates to a method for predicting functions of a protein based on amino acid sequence information, where the protein is constituted by the said amino acid sequence. More specifically, the invention relates to a method for effectively predicting, based on amino acid sequence information, biological functions of a protein constituted by said amino acid sequence, such as enzyme activities, by using a specific database available for computer.

Background Art

Proteins are essential substances for the maintenance of life activities in organisms, and various proteins exist in animals and plants as well as microorganisms, which bear characteristic functions and roles. Proteins can be roughly categorized with their functions into enzymes which catalyze chemical reactions, receptors which are receptor proteins of signal transducing substances, signal transducing proteins which transfer signals themselves, and proteins which bind and transport specific substances, and each can be further subdivided by various functions. For example, enzymes catalyze specific reactions such as enzymes which reduce specific parts of particular substrates and enzymes which hydrolyze proteins.

Proteins are mainly composed of 20 kinds of amino acids, and are polypeptide molecules in which 50 to 1000 amino acids are linked as a chain in various orders by polypeptide bonds. The order of amino acid linkage (called amino acid sequence or primary structure) is different for each protein, and as a result, each protein can exhibit different physiological functions. That is, once a long polypeptide chain is folded into a certain tree-dimensional structure, capture of target molecules (enzyme substrate molecules, receptor substrate molecules and the like) becomes possible and functional groups related to a reaction are placed at appropriate positions, providing a suitable field for appearance of the target biological function. It is easily predicted that a characteristic steric structure is determined from each amino acid sequence and biological functions are determined from the steric structure, however, the inevitability of those relationships is not well explained.

For the study of proteins, methods of determining a whole amino acid sequence

from a gene by picking up the gene encoding a protein, after determining the terminal 20 or less residues and synthesizing the corresponding gene sequence, are being utilized instead of the classical methods, in which a protein is isolated and purified using enzyme activity as an index, the molecular weight, constituting amino acid numbers, and numbers of each amino acid are determined, and finally an amino acid sequence is determined. In these studies, proteins with known functions have been on target, however recently, completely reversed orders of studies are conducted in many cases. The reason is the fact that analysis of gene sequences has become quite easy, and consequently, it has become possible to determine an amino acid sequence of a protein from the gene without isolating the protein.

Consequently, proteins are rapidly increasing for which only amino acid sequences are predicted and whose biological functions remain unknown. Since the biological function of a protein appears based on the steric structure, trials have been made to predict biological functions of such proteins with unknown functions by analyzing the steric structures using crystallographic analysis and NMR analysis. However, for these structural analyses, much larger amount and also highly purified samples are required compared with biochemistry works. Since biological functions are not always predicted from the steric structures and even predicted biological functions are not necessarily important ones, there is also a problem that investment efficiencies of such studies are awfully poor. Therefore, it is desired earnestly to develop a method of predicting biological functions of a protein possessing a certain amino acid sequence before determining its steric structure. If such a method is developed, it will be expected to contribute a great deal to the protein study and genomic study.

Steric structures and biological functions are closely related, and thus steric structure information on proteins with known functions is useful for various purposes, not only for elucidating mechanism of the functions. Three-dimensional coordinates of proteins or complexes with ligand molecules are stored in the Protein Data Bank (Brookhaven National Laboratories, U.S.A.), which is accessible throughout the world. At present, the number of the structures stored is about 5,000, but considering independent proteins excluding the difference in species and mutants, the number becomes about 400 to 500. Although the number of proteins that are analyzed crystallographically is increasing with accelerating speed due to the propagation of

analysis techniques and the progress of isolation and purification techniques of proteins, proteins whose structures are not yet solved are overwhelming majority at present.

It is possible to utilize modeling as a means to predict steric structures of proteins and modes of interactions with ligand molecules, besides crystallographic analysis and NMR analysis. When steric structures of similar proteins with moderately high homology of amino acid sequences have already been analyzed, a steric structure based on the correspondence of amino acid residues can be constructed by performing modeling using the structures as templates. This method has advantages that there is no need to obtain samples, and that the method is generally performed interactively on the computer graphics screen. For instance, for those amino acids that are not identical, it is conducted by replacing the side chains. Although there are problems such as the conformation of side chains and the insertion or deletion of amino acids of main chains, reliability of the predicted structures is governed by the degree of homology of the amino acid sequences and it is possible to treat it like crystal structures almost similarly.

A method of making a correspondence between amino acid sequences of two or more kinds of proteins (alignment) such that the kinds of amino acids match as many as possible, is frequently used for the purpose of examining similarity and difference between species and families as well as for modeling. In the conceptual technique of alignment, a corresponding position with the best coincidence score is found while the sequence is subjected to sliding one by one relative to the other. Actually, however, precise consideration and repeated operations are necessary, because possible correspondences between the sequences are unlimited, and thus extremely complicated operations are necessary in order to obtain an accurate result. For example, the degree of coincidence of amino acids with whole sequences can not simply be taken as a score, because there are often insertions or deletions in one sequence, and thus it is necessary to find a partial sequence which coincides well locally, and also it is necessary to calculate coincidence scores in a unit of certain number of residues (window). In some cases, it is necessary to calculate scores by regarding not strictly identical but similar amino acids as homologous.

Generally, when the homology is low, there is a problem that an alignment is not determined unanimously and still has some uncertainty. However, at present,

alignment is the simplest and easiest method of estimating the functions of a protein with the amino acid sequence only as a clue, since proteins with high similarity of the amino acid sequences can be found out from groups of proteins whose functions are known. For these reasons, trials of partially automating the complicated operations of alignment by using a computer have been made. For example, as matching methods of amino acid sequences, FASTA (Pearson W. R. and Lipman, D. J., Proc. Natl. Acad. Sci. USA, 85, pp.2444-244, 1988) and BLAST (Altschul, S. F. et al., J. Mol. Biol., 215, pp.403-410, 1990) are known. These methods are suitable for rapid examination of the existence of a short specific sequence in a long target sequence. However, when long sequences are compared with each other or when sequences with low homology and fragmental coincidence are treated as search targets, judgment of similarity and extraction of similar parts are extremely difficult and accuracy of judging the homology is low. Therefore, these methods are not adequate for the purpose of alignment and prediction of protein functions, and thus a development of a rapid method with better accuracy has been desired.

Disclosure of Invention

An object of the present invention is to provide a method for predicting functions of a protein based on amino acid sequence information, where the protein is constituted by the said amino acid sequence. More specifically, the object is to provide, when information on amino acid sequences is only available, a method for effectively searching homology to amino acid sequences of proteins with known biological functions with a computer using a specific database, and then searching biological functions of the protein constituted by said amino acid sequence accurately and rapidly.

As a result of the inventors' earnest effort to solve the above-mentioned object, it was discovered that protein functions can be predicted extremely rapidly and accurately from an amino acid sequence when the database with the following characteristics is employed.

The present invention thus provides a database comprising information on amino acid sequences of proteins with one or more known biological functions, and further comprising information on importance scores concerning the appearance of said biological functions which is added to each amino acid residue constituting said

amino acid sequences.

This database, for instance, can be used to predict functions of a protein with unknown biological functions based on the homology of amino acid sequences. According to a preferable embodiment, as information on the amino acid sequences of proteins with known biological functions, importance scores concerning the binding between the proteins and ligand molecules or concerning the appearance of biological functions are added to each amino acid residue constituting the said amino acid sequences, by using the amino acid sequences of proteins for which information on the steric structure such as the three dimensional structures of protein is available. These databases, in general, can be stored in various media such as floppy disks, CD-ROM, magnetic tapes, and optical disks.

From other point of view, the present invention provides a method of preparing an alignment of a protein in the above-mentioned database (referred to as "template protein" in the specification) and a polypeptide with unknown biological functions (referred to as "target protein" in the specification), which comprises the steps of calculating homology measure for the coincidence of constituting amino acids under consideration of the importance scores concerning the appearance of biological functions, and preparing an alignment which represents the homology of regions where the said importance is high.

A preferred embodiment of the aforementioned method comprises a step of searching correspondences with high homology regarding the protein in the above-mentioned database and the target protein, by using group sequences containing two or more continuous amino acid residues that are highly important for the appearance of the biological functions. Furthermore, other preferred embodiment includes a method comprises a step of obtaining a final score of homology from the above-mentioned alignment for one of the proteins in the database and the target protein, and a method comprises a step of estimating one or more proteins most similar to the target protein with regard to the biological functions based on the final scores for all the proteins stored in the database. These methods have characteristics that, even if homology of the whole protein is low, proteins having high homology in regions related to the biological functions can be extracted and functions of the target protein can be rapidly and very accurately estimated.

Brief Explanation of Drawings

Figure 1 shows information on amino acid sequences with the importance scores concerning the appearance of biological functions added for each amino acid residue constituting the amino acid sequence, for 4 kinds of proteins with known biological functions and steric structures. Symbols in the figure indicate dihydrofolate reductase from E. coli (DHFR-EC), trypsin from bovine (TRYP), ribonuclease from bovine (RNAS), and myoglobin from whale (MYGL), and amino acid residues are indicated by one-letter symbols.

Figure 2 shows an alignment of the target protein (DHFR-HM) and DHFR-EC that was extracted as a template protein yielding the highest SSS score to the target protein. The symbols in the figure indicate dihydrofolate reductase from human (DHFR-HM), and dihydrofolate reductase from E. coli (DHFR-EC) respectively, and first row indicates the amino acid numbers of DHFR-HM, second row indicates the amino acid sequence of DHFR-HM, third row indicates the partial sequences from the amino acid sequence of DHFR-EC, and fourth row indicates the amino acid numbers of partial amino acid sequence of DHFR-EC.

Best Mode for Carrying out the Invention

Database of the present invention is characterized that it includes information on amino acid sequences of proteins with one or more known biological functions, and that it includes information on importance scores concerning the appearance of biological functions added to each amino acid residue constituting the said amino acid sequences. Proteins to be stored can be, for example, any protein insofar as that one or more biological functions like enzyme actions and receptor actions are known and that all amino acid sequences are known. Proteins whose steric structures of proteins or ligand-binding sites have already been elucidated or predicted, or easily predictable, are preferable. In the database, it is desirable to contain information on the protein with known steric structures as much as possible. For instance, three-dimensional coordinates of proteins or complexes with ligand molecules are stored in the Protein Data Bank (Brookhaven National Laboratories, U.S.A.), where information on about 5,000 proteins (about 400 to 500 considering independent proteins excluding the difference in species and mutants) is available, and can be used suitably to make a database of the present invention.

In the database of present invention, it is characterized that, based on the amino acid sequences constituting proteins with known biological functions, importance scores concerning the appearance of the said biological functions are added to each amino acid. As information to be stored for each protein, examples include, for example, name of the protein, species, organ and apparatus, subtype, kind of function, sub-classification of the function (for example, for enzymes, enzyme action such as protein degradation and reduction), enzyme classification number (EC number), ligand molecules related to the enzyme reactions or biological functions (enzyme substrate, receptor substrate, coenzyme, metal ion, and the like), source of the steric structure (for example, X-ray crystallography, NMR analysis, modeling based on information on similar proteins with similar biological functions), main bibliographic references, reference number of other databases, target region ligands, and whole amino acid sequences. However, information is not limited to these examples, and some information may be added or deleted appropriately.

In addition to the above-mentioned information, the database of the present invention includes information on importance scores concerning the appearance of the said biological functions, which is added to each amino acid residue constituting the said amino acid sequences. Importance scores are specified by giving numbers or other symbols, for example, 0 for no importance and 10 for extremely high importance. Preferably, it is general to calculate the scores considering two or more elements related to the importance. To the database of the present invention, it is further possible to add information such as existence of continuous amino acid residue sequence (1 to n) that contributes to the appearance of the biological functions (i.e. score is not zero), method of the scoring, sum of the scores, scale factors for normalizing the sum of the scores among proteins. However, information to be added is not limited to these examples, and appropriate information may be added or deleted. When multiple biological functions or multiple ligand molecules are known for one protein, it is preferable to store information for each, respectively.

In the following descriptions, procedures will be demonstrated for giving importance scores concerning the appearance of biological functions for each amino acid residue constituting amino acid sequences of proteins. The descriptions are presented only as examples and should not be interpreted in any limiting sense. Furthermore, the database of the present invention is not limited to those produced by

these procedures. In the following descriptions, examples of indicating the importance scores by numerical values will be explained, in which zero is given to those without any importance and greater numbers are given to those with increasing importance. However, it should be recognized that the scoring method is not limited to these examples.

(a) In case crystallographic analysis of protein complexes including ligand molecules of small molecular weights has been conducted:

For proteins stored in the Protein Data Bank whose three-dimensional structures in complex with small-molecular-weight ligand molecules such as an enzyme substrate, a receptor substrate or an inhibitor have been analyzed, a distance of each amino acid residue from the ligand molecule can be calculated and importance scores depending on the distance of each amino acid residue are given. As the small-molecular-weight ligand molecules, any molecules may be employed such as, for example, an organic compound that is pharmacologically active, an enzyme substrate, and a metal ion. For instance, a score of 1 can be given to any amino acid residue which is within 10 Å from any atom of the ligand molecule (for example, Ca atom which becomes a ligand or one atom in the side chain of the ligand molecule), a score of 2 if it is within 8 Å, 4 if it is within 6 Å, and 0 if it is more than 10 Å.

(b) In case crystallographic analysis of unbound protein has been conducted:

For proteins for which crystallographic analyses have been performed without ligand molecules, if structural region related to biological functions (such as enzyme activities) can be predicted from various experiments, it is possible to give numerical values to the neighboring amino acid residues depending on the distance, as is performed in above (a). Even if the correspondence between the biological function and steric structure is not established, insofar as the biological functions are clear, amino acid residues related to the biological functions can be extracted by searching distinguished cavities while the drawing of steric structure of the protein is being rotated on computer graphics screen.

(c) Structurally conserved region:

For proteins having the same biological functions, if analytical results can be utilized as for the steric structures of two or more kinds such as subtypes with different amino acid sequences and proteins from different species, a structurally conserved region can be extracted by superposing those structures, and high numerical

values can be given to the amino acid residues contained in those regions.

(d) Modeling:

In case the steric structures of proteins have not been analyzed, it is possible to give importance scores based on the modeling structure constructed on the basis of the steric structures of homologous proteins that are known to have practically the same biological functions. It is known that reliability of the modeling structure is high, for example, for those cases such as receptor subtypes, isozymes, proteins of the same family, proteins of different species with the same functions whose amino acid sequences have high homology. The method of giving importance scores, for example, may be conducted similarly to the above-mentioned techniques.

(e) Biochemical experiment and genomic experiment:

High importance scores can be given to the amino acid residues that are predicted to be important for the appearance of the biological functions from biochemical experiments and the like, and to the amino acid residues that are predicted to be essential for the appearance of the biological functions such as enzyme reactions from the genomic amino acid conversion experiment (point mutation and the like). In enzyme reactions, for example, large numerical values can be given to amino acid residues that play catalytic role, in addition to the evaluation from the bonding with ligand molecules.

(f) Protein as macromolecular ligand molecule

Generally, proteins in which binding with a small ligand molecule is essential for its function have cavities that bind stably with the ligand molecules. On the other hand, for proteins to which macromolecular ligand molecules like proteins bind, it is frequently observed that they bind with receptor proteins through the molecular surface without forming distinguished cavities, and there are cases where the receptor proteins do not have distinguished cavities. For example, in the case of cytokines which become macromolecular ligand molecules themselves, large numerical values may be given to the amino acid residues in the epitope region which are predicted by using monoclonal antibodies.

By employing a combination of one or more kinds of techniques exemplified above, and further by adding appropriate techniques, if required, importance scores concerning the appearance of the said biological functions can be given to each amino acid residue constituting the amino acid sequences of proteins with known functions,

and information on the amino acid sequences attached to the importance scores can be prepared. As for the relationship between amino acid residues and appearance of biological functions, it should be understood that various criteria can be utilized for scoring such as those whose relationship can be predicted to some extent, as well as those whose relationship has been fully proved, for example, by the above-mentioned method (a).

For instance, information can be collected on the amino acid sequence attached to the importance scores for as many proteins as possible such as those with known crystal structures and those predicted to have similar steric structures from biological functional points of view, and then a database of the present invention can be constructed by storing the information in a certain format usable by computers. For this purpose, depending on the quantity and quality of the information on each protein, scores may be given to each protein based on appropriate different criteria. However, there are some cases where the scoring method and a scaling factor used for normalization of the total scores among the proteins need to be added to the database. While it is possible to input the above-mentioned information manually according to a certain method, it is generally effective to perform it by using a certain program on computer graphics screen.

According to the method of the present invention, using the above-mentioned database, an alignment can be prepared concerning the template protein whose information is stored in the database and a target protein with unknown biological functions in such a manner that the homology score calculated from the importance scores of amino acid residues becomes maximum, then similar alignments can be prepared for more than 2 template proteins, or preferably all template proteins, in the database. Subsequently, a template protein with the highest score is selected by comparing homology scores among the template proteins. It can be estimated that the template protein thus selected has high similarity of steric structures to the target protein and has practically the same biological functions.

Above method, in general, is performed by taking out information on the template proteins in the above-mentioned database of the present invention one by one, and by performing an alignment to the amino acid sequence of the target protein. If the information on the amino acid sequence of the target protein is directly available, that information may be input and utilized, and if only the information on the genomic

sequence encoding the target proteins is available, it is necessary to use information on the amino acid sequence interpreted from the information of its nucleic acid sequence.

As an example of a preferable method of the present invention includes a method of searching a correspondence with high homology concerning the template proteins and the target protein, by using group sequences comprising 2 or more continuous amino acid residues in the amino acid sequence of template proteins (continuous amino acid residues with scores other than zero) that contribute to the biological functions. However, the alignment procedure is not limited to this method, and may be performed by any method available for those skilled in the art.

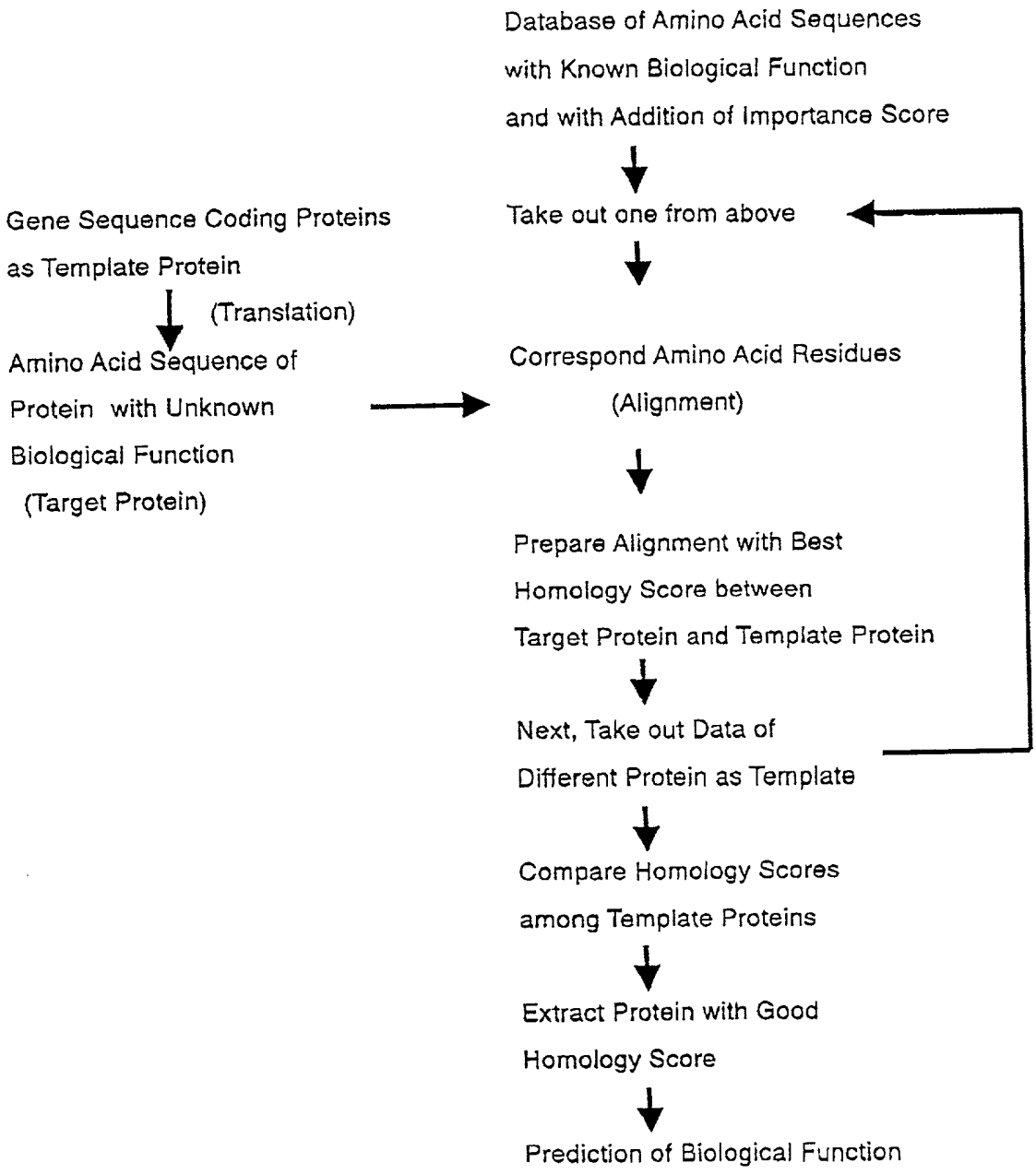
By the above-mentioned method which employs group sequences, a homology score for each group is obtained by sliding the group sequence one by one against the amino acid sequence of the target protein, and thereafter, a correspondence of each group sequence to the amino acid sequence of the target protein is determined in such a manner that a total score of all group sequences becomes best while considering factors such as the linkage order and lengths of the group sequences as necessary. This procedure can be performed for all template proteins in the database, and one or more proteins having high total scores can be extracted. It is highly possible that the target protein possesses practically the same biological functions as the template proteins thus extracted.

As for the score of homology, for example, if a group sequence in the template protein coincides with the corresponding amino acid residue in the target protein sequence, it is easy to transfer the importance score to the coincided amino acid residue and simply sum them up. However, in order to prepare an alignment emphasizing the coincidence of amino acids with high importance scores, each importance score may be processed further by one or more functions and then used. Upon preparation of alignments between the target protein and all template proteins contained in the database, total homology scores among the different alignments may be compared. Generally, however, it is desirable to calculate scale factors for normalization of importance scores for each template protein and store them in the database so that goodness of homology can be compared by the sum of the importance scores even for template proteins of different sizes. At the step where alignment between the target protein and each template protein is completed and homology

scores are calculated, final scores can be obtained by multiplying the homology score of each template protein by the corresponding scale factor, and goodness or badness of homology among the template proteins can be determined.

In the same kind of proteins existing in different species, for example, aspartic acid and glutamic acid, that have carboxyl group in common and whose side-chain lengths differ only by one carbon atom, sometimes play the same role on similar positions in amino acid sequences. In such cases, it is appropriate to regard these amino acid residues as being coincided, upon judgments of the coincidence of amino acid residues. Furthermore, although amino acid residues such as leucine, isoleucine, and valine differ in the shape and size (bulkiness), they have similar characteristics from a viewpoint of hydrophobicity. Therefore, upon evaluation of the homology of amino acid sequences, it is desirable to use correspondence table that grades the similarity of amino acid residues in order to reflect the existence of these analogous amino acid residues. Although any measures may be used for the similarity of amino acid residues, PAM250 (Dayhoff, M.O., et al., Atlas of Protein Sequence and Structure, Dayhoff, M. O. Ed., Vol. 5, Suppl. 3, pp.345-352, NBRF, Washington, 1978) and BLOSUM (Henikoff, S. and Henikoff, J. G., Proc. Natl. Acad. Sci. USA, 89, pp.10915-10919, 1992) are available examples as correspondence tables describing the similarity.

Table 1



An example of the method of the present invention is shown above as a schematic chart. For example, the method of the present invention may be a method including the following processes. However, a method of the present invention is not limited to these methods, and it should be understood that one or two or more appropriate processes may be added, if necessary, in addition to the processes employed in these

methods, and that there are occasions in which one or two or more processes may be omitted, if desired. It should be also noted that such modified or altered methods are all included within the scope of the present invention.

(1) Process of obtaining amino acid sequence of a target protein;
(2) Process of selecting one template sequence from the above-mentioned database;
(3) Process of taking out partial sequences *a, b, c, d, e, ... ,n, ...*, with importance scores greater than a certain value, from amino acid sequence of the template protein, for example, from N-terminal successively (length of each partial sequence is represented as *la, lb, lc, ld, le, ..., ln, ...*);
(4) Process of placing the partial sequence *a* at the first position of the target sequence and calculating homology score *S(a)i* while sliding the amino acid residues one by one (add importance scores of amino acid residues coincided as a homology score);
(5) Process of placing the partial sequence *b* at the (*l+la*)-th position of the target sequence and calculating homology score *S(b)i* while sliding the amino acid residues one by one (add importance score of amino acid residues coincided as a homology score);
(6) Processes of calculating homology scores *S(n)i* similarly for *c, d, e, ..., n, ...*;
(7) Process of determining corresponding positions of partial sequences so that homology SS of all partial sequences become the highest, considering the order *a, b, c, d, e, ..., n, ...* and the numbers of amino acid residue of each;
(8) Process of multiplying SS by scale factors to give SSS;
(9) Process of obtaining SSS by the above procedures for all template proteins in the database; and
(10) Process of extracting proteins with high SSS values.

Example

The present invention is explained more specifically by referring to the following examples. However, the scope of the present invention is not limited to the following examples.

Example 1: Preparation of database

A database was prepared which comprised four kinds of proteins with known

biological functions and known steric structures which are dihydrofolate reductase from E. coli (DHFR-EC), trypsin from bovine (TRYP), ribonuclease from bovine (RNAS), and myoglobin from whale (MYGL). Each crystal structure was obtained from the Protein Data Bank (Brookhaven National Laboratories, USA). Information was prepared on the amino acid sequences with importance scores concerning the appearance of each biological function added to each amino acid residue constituting each amino acid sequence, giving a score of 2 when any constituting atom of an amino acid residue is located within 4 Å from any atom of the ligand molecule (inhibitor or coenzyme), score of 1 for the range between 4 to 10 Å, score of 0 for others. Figure 1 shows the amino acid sequence of each protein and the results of the scoring.

Example 2: Prediction of biological functions

Biological functions of a target protein was predicted by the method of the present invention, by taking dihydrofolate reductase from human (DHFR-HM) as the target protein and using the above-mentioned database. Although DHFR-HM is a protein with known biological functions and known steric structure, the analysis was performed under assumption that the biological functions and steric structure were unknown. Partial sequences with a score of 1 or more for each amino acid sequence of the template proteins in the database was taken out, and homology score S was calculated by sliding residue one by one against the amino acid sequence of the target protein, and an alignment position with the highest value of S score was determined.

For the calculation of S score, the correspondence table BLOSUM62 was used which concerns similarity of amino acids, and a method was employed in which homology scores was obtained from the table, which corresponded to sets of amino acid residues between the partial sequence and the amino acid sequence of the target protein, and a summation of the product of the score for each residue of the partial sequence and the similarity scores was taken over the length of the partial sequence. Homology SS of all partial sequences was calculated by obtaining the largest value of S scores for each partial sequence and by taking a summation of those values. In order to compensate for the difference in lengths of the sequences used in the homology determination, reciprocal numbers of sums of scores for all partial sequences as scale factors was used, and final scores SSS was calculated by multiplying SS by the scale factors. Consequently, as shown in Table 1, when DHFR-EC, TRYP, RNAS, MYGL

were employed as template proteins, DHFR-EC gave the highest score SSS, and it was predicted that the target protein (DHFR-HM) is similar to DHFR-EC and has activities of dihydrofolate reductase. Figure 2 shows the alignments of DHFR-HM and DHFR-EC.

Table 2

Protein	SSS Score
DHFR-EC	1.82
TRYP	1.09
RNAS	1.22
MYGL	0.61

Industrial Application

The database of the present invention is useful for prediction, based on the information on amino acid sequences, of biological functions of proteins constituted of those amino acid sequences. The method of the present invention is useful for accurate and rapid search of biological functions of proteins comprising the amino acid sequences using said database.

What is claimed is:

1. A database comprising information on amino acid sequences of proteins with one or more known biological functions, and further comprising information on importance scores regarding appearance of said biological functions added for each amino acid residue constituting said amino acid sequences.

2. The database according to claim 1 which is utilized to predict function of a protein with unknown biological function based on homology of amino acid sequences.

3. The database according to claim 1 or claim 2 which is prepared by using amino acid sequences of proteins for which information on protein steric structures is available as the information on amino acid sequences of proteins with known biological functions.

4. The database according to any one of claims 1 through 3 which is stored in a storage medium.

5. A method of preparing an alignment of a protein stored in the database according to any one of claims 1 through 3 and a polypeptide with unknown biological function which comprises the steps of:

calculating a homology score to the coincidence of each constituent amino acid under consideration of the importance score for the appearance of a biological function, and

preparing an alignment representing homology of sites where said importance is high.

6. The method according to claim 5, which comprises the step of searching correspondence with high homology concerning the protein in the database and the target protein by using group sequences comprising two or more continuous amino acid residues of high importance with regard to the appearance of the biological function.

7. The method according to claim 5 or claim 6, which comprises the step of obtaining a final score of homology from the alignment regarding one of the proteins in the database and the target protein.

8. The method according to claim 7, which comprises the step of predicting one or more proteins most homologous to the target protein with regard to the biological function based on the final scores for all the proteins in the database.

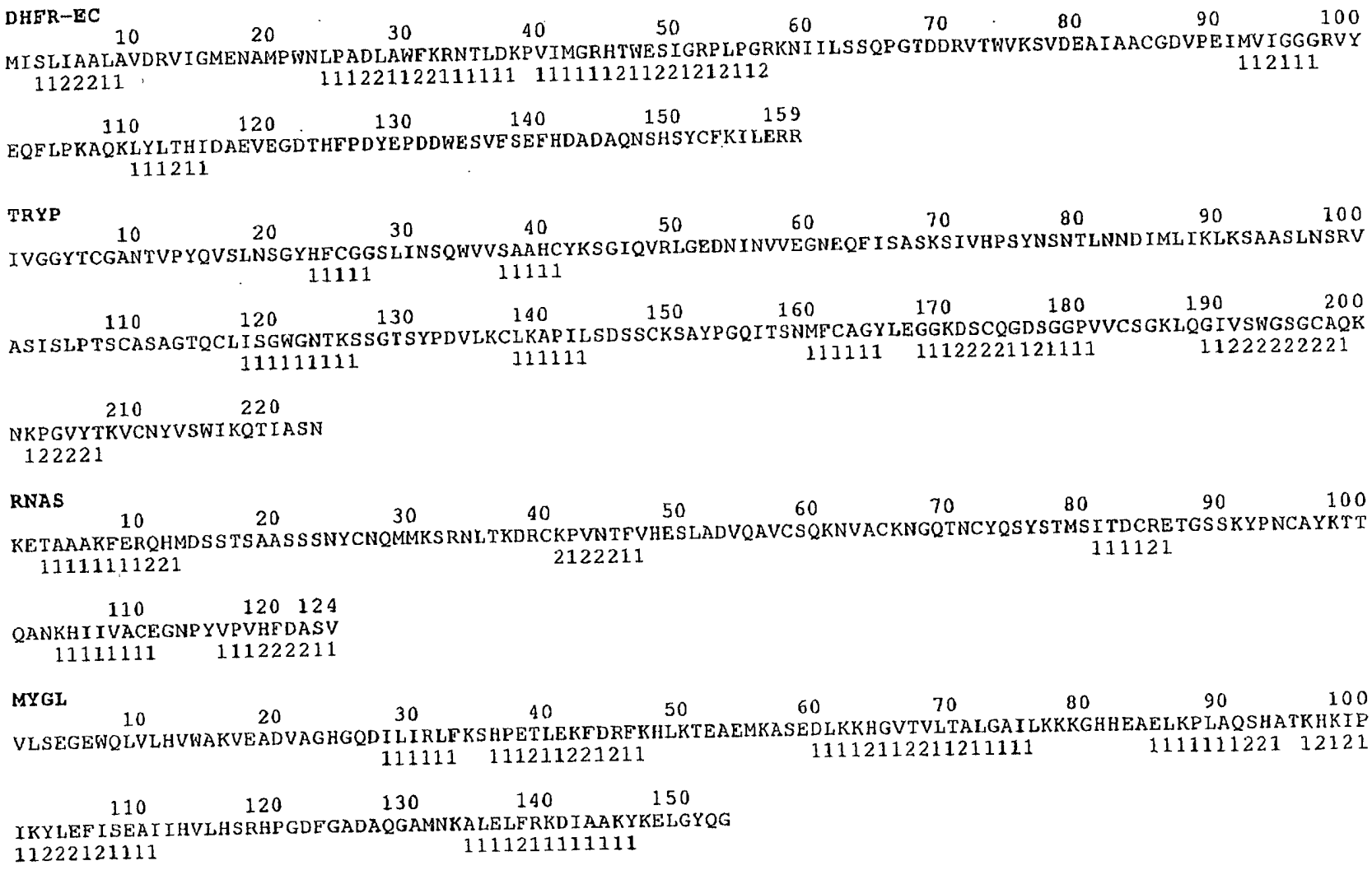

DHFR-EC
10 20 30 40 50 60 70 80 90 100
MISLIAALAVDRVIGMENAMPWNLPADLAWFKRNTLDKPVIMGRHTWESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVIGGGRVY
1122211 111221122111111 111111211221212112 112111
110 120 130 140 150 159
EQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFHDADAQNSHSYCFKILERR
111211
TRYP
10 20 30 40 50 60 70 80 90 100
IVGGYTCGANTVPYQVSLNSGYHFCGGSLINSQWVVSAAHCYKSGIQVRLGEDNINVVEGNEQFISASKSIVHPSYNSNTLNNDIMLIKLKSAASLNSRV
11111 11111
110 120 130 140 150 160 170 180 190 200
ASISLPTSCASAGTQCLISGWGNTKSSGTSYPDVLKCLKAPILSDSSCKSAYPGQITSNMFCAGYLEGGKDSCQGDSGGPVVCSGKLQGIVSWGSGCAQK
111111111 111111 111111 11122221121111 11222222221
210 220
NKPGVYTKVCNYVSWIKQTIASN
122221
RNAS
10 20 30 40 50 60 70 80 90 100
KETAAAKFERQHMDSSTSAASSSNYCNQMMKSRNLTKDRCKPVNTFVHESLADVQAVCSQKNVACKNGQTNCYQSYSTMSITDCRETGSSKYPNCAYKTT
11111111221 2122211 111121
110 120 124
QANKHIIVACEGNPYVPVHFDASV
11111111 111222211
MYGL
10 20 30 40 50 60 70 80 90 100
VLSEGEWQLVLHVWAKVEADVAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEMKASEDLKKHGVTVLTALGAILKKKGHHEAELKPLAQSHATKHKIP
111111 111211221211 11112112211211111 1111111221 12121
110 120 130 140 150
IKYLEFISEAIIHVLHSRHPGDFGADAQGAMNKALELFRKDIAAKYKELGYQG
11222121111 1111211111111

FIG. 1

1/2

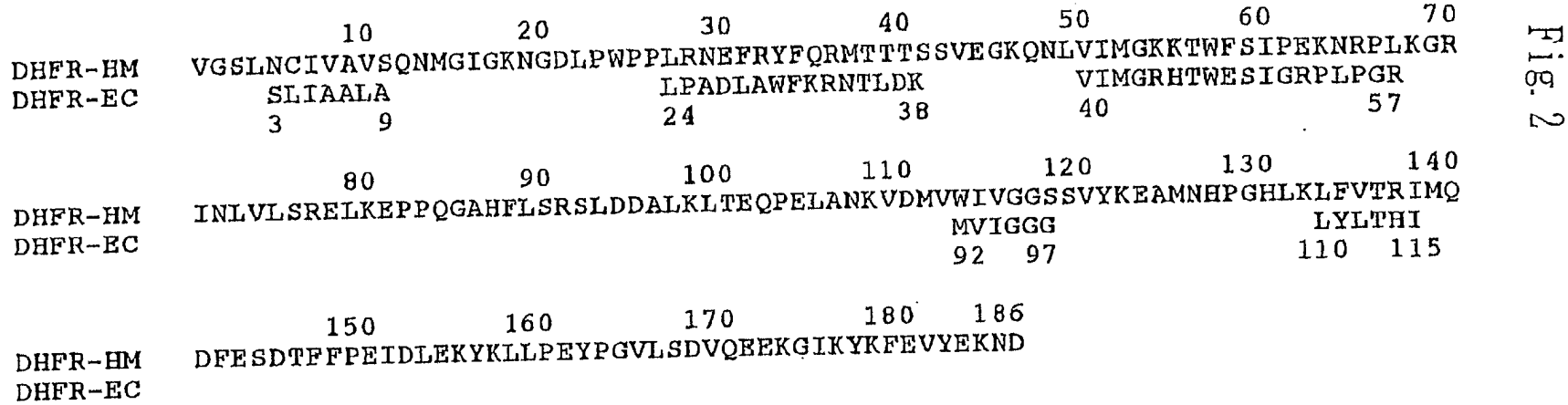


FIG. 2

SEQUENCE LISTING

<110> ITAI, Akiko
ITAI, Reiko
TOMIOKA, Nobuo

<120> Method For Predicting Functions of Protein

<130> P20294

<140> 09/700,708
<141> 2000-11-24

<150> PCT/JP98/02302
<151> 1998-05-26

<160> 6

<170> PatentIn version 3.0

<210> 1
<211> 159
<212> PRT
<213> Escherichia coli

<400> 1

Met Ile Ser Leu Ile Ala Ala Leu Ala Val Asp Arg Val Ile Gly Met
1 5 10 15

Glu Asn Ala Met Pro Trp Asn Leu Pro Ala Asp Leu Ala Trp Phe Lys
20 25 30

Arg Asn Thr Leu Asp Lys Pro Val Ile Met Gly Arg His Thr Trp Glu
35 40 45

Ser Ile Gly Arg Pro Leu Pro Gly Arg Lys Asn Ile Ile Leu Ser Ser
50 55 60

Gln Pro Gly Thr Asp Asp Arg Val Thr Trp Val Lys Ser Val Asp Glu
65 70 75 80

Ala Ile Ala Ala Cys Gly Asp Val Pro Glu Ile Met Val Ile Gly Gly
85 90 95

Gly Arg Val Tyr Glu Gln Phe Leu Pro Lys Ala Gln Lys Leu Tyr Leu
100 105 110

Thr His Ile Asp Ala Glu Val Glu Gly Asp Thr His Phe Pro Asp Tyr
115 120 125

Glu Pro Asp Asp Trp Glu Ser Val Phe Ser Glu Phe His Asp Ala Asp
130 135 140

Ala Gln Asn Ser His Ser Tyr Cys Phe Lys Ile Leu Glu Arg Arg
145 150 155

<210> 2
<211> 223
<212> PRT
<213> Bovine

<400> 2

Ile Val Gly Gly Tyr Thr Cys Gly Ala Asn Thr Val Pro Tyr Gln Asx
1 5 10 15

Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Ser
20 25 30

Gln Trp Val Val Ser Ala Ala His Cys Tyr Lys Ser Gly Ile Gln Val
35 40 45

Arg Leu Gly Glu Asp Asn Ile Asn Val Asx Glu Gly Asn Glu Gln Phe
50 55 60

Ile Ser Ala Ser Lys Ser Ile Val His Pro Ser Tyr Asn Ser Asn Thr
65 70 75 80

Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Lys Ser Ala Ala Ser Leu
85 90 95

Asn Ser Arg Val Ala Ser Ile Ser Leu Pro Thr Ser Cys Ala Ser Ala
100 105 110

Gly Thr Gln Cys Leu Ile Ser Gly Trp Gly Met Thr Lys Ser Ser Gly
115 120 125

Thr Ser Tyr Pro Asp Asx Leu Lys Cys Leu Lys Ala Pro Ile Leu Ser
130 135 140

Asp Ser Ser Cys Lys Ser Ala Tyr Pro Gly Gln Ile Thr Ser Asn Met
145 150 155 160

Phe Cys Ala Gly Tyr Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp
165 170 175

Cys Gly Gly Pro Val Val Cys Ser Gly Lys Leu Gln Gly Ile Val Ser
180 185 190

Trp Gly Ser Gly Cys Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys
195 200 205

Val Cys Asn Tyr Val Ser Trp Ile Lys Gln Thr Ile Ala Ser Asn
210 215 220

<210> 3
<211> 124
<212> PRT
<213> Bovine

<400> 3

Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser Ser
1 5 10 15

Thr Ser Ala Ala Ser Ser Ser Asn Tyr Cys Asn Gln Met Met Lys Ser
20 25 30

Arg Asn Leu Thr Lys Asp Arg Cys Lys Pro Val Asn Thr Phe Val His
35 40 45

Glu Ser Leu Ala Asp Val Gln Ala Val Cys Ser Gln Lys Asn Val Ala
50 55 60

Cys Lys Asn Gly Gln Thr Asn Cys Tyr Gln Ser Tyr Ser Thr Met Ser
65 70 75 80

Ile Thr Asp Cys Arg Glu Thr Gly Ser Ser Lys Tyr Pro Asn Cys Ala
85 90 95

Tyr Lys Thr Thr Gln Ala Asn Lys His Ile Ile Val Ala Cys Glu Gly
100 105 110

Asn Pro Tyr Val Pro Val His Phe Asp Ala Ser Val
115 120

<210> 4
<211> 153
<212> PRT
<213> Whale

<400> 4

Val Leu Ser Glu Gly Glu Trp Gln Leu Val Leu His Val Trp Ala Lys
1   5   10   15

Val Glu Ala Asp Val Ala Gly His Gly Gln Asp Ile Leu Ile Arg Leu
            20   25   30

Phe Lys Ser His Pro Glu Thr Leu Glu Lys Phe Asp Arg Phe Lys His
        35   40   45

Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys His
    50   55   60

Gly Val Thr Val Leu Thr Ala Leu Gly Ala Ile Leu Lys Lys Lys Gly
65   70   75   80

His His Glu Ala Glu Leu Lys Pro Leu Ala Gln Ser His Ala Thr Lys
                85   90   95

His Lys Ile Pro Ile Lys Tyr Leu Glu Phe Ile Ser Glu Ala Ile Ile
            100   105   110

His Val Leu His Ser Arg His Pro Gly Asp Phe Gly Ala Asp Ala Gln
        115   120   125

Gly Ala Met Asn Lys Ala Leu Glu Leu Phe Arg Lys Asp Ile Ala Ala
    130   135   140

Lys Tyr Lys Glu Leu Gly Tyr Gln Gly
145   150

<210> 5
<211> 186
<212> PRT
<213> Human

<400> 5

Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly Ile
1   5   10   15

Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Arg Asn Glu Phe Arg
            20   25   30

Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln Asn
        35   40   45

Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile Pro Glu Lys Asn
    50   55   60

```
Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
65                  70                  75                  80

Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
                85                  90                  95

Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
            100                 105                 110

Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala Met Asn His Pro
        115                 120                 125

Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln Asp Phe Glu Ser
    130                 135                 140

Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
145                 150                 155                 160

Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
                165                 170                 175

Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
            180                 185
```

<210> 6
<211> 186
<212> PRT
<213> Escherichia coli

<400> 6

```
Val Gly Ser Leu Ser Leu Ile Ala Ala Leu Ala Gln Asn Met Gly Ile
1               5                   10                  15

Gly Lys Asn Gly Asp Leu Pro Trp Pro Pro Leu Pro Ala Asp Leu Ala
            20                  25                  30

Trp Phe Lys Arg Asn Thr Leu Asp Lys Ser Val Glu Gly Lys Gln Asn
        35                  40                  45

Leu Val Ile Met Gly Arg His Thr Trp Glu Ser Ile Gly Arg Pro Leu
    50                  55                  60

Pro Gly Arg Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
65                  70                  75                  80

Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
```

```
                    85                  90                  95

Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
            100                 105                 110

Met Val Ile Gly Gly Gly Ser Val Tyr Lys Glu Ala Met Asn His Pro
        115                 120                 125

Gly His Leu Lys Leu Tyr Leu Thr His Ile Met Gln Asp Phe Glu Ser
    130                 135                 140

Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
145                 150                 155                 160

Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
                165                 170                 175

Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
            180                 185
```

Case # P20294

DA 0537M/US

#3

Declaration and Power of Attorney For Utility or Design Patent Application
特許出願宣言書

Japanese Language Declaration

私は、下欄に氏名を記載した発明者として、以下のとおり宣言する：

私の住所、郵便の宛先および国籍は、下欄に氏名に続いて記載したとおりであり、

名称の発明に関し、請求の範囲に記載した特許を求める主題の本来の、最初にして唯一の発明者である(一人の氏名のみが下欄に記載されている場合)か、もしくは本来の、最初にして共同の発明者である(複数の氏名が下欄に記載されている場合)と信じ、

_____

_____

_____

☐ 上記発明の明細書(下記の欄でX印がついていない場合は、本書に添付)は、

☐ _____年_____月_____日に提出され、

米国出願番号 _____ とし、

(該当する場合) _____年_____月_____日に訂正されました。又は、

☐ 特許協定条約国際出願番号 _____ とし、

(該当する場合) _____年_____月_____日に訂正されました。

私は、前記のとおり補正した請求の範囲を含む前記明細書の内容を検討し、理解したことを陳述する。

私は、連邦規則法典第37編第1条第56項に定義されるとおり、特許資格の有無について重要な情報を開示すべき義務があることを認めます。

私は合衆国法典第35部第119条(a-d)項又は第365条(b)項に基づく、下記の外国特許出願又は発明者証出願、或いは第365条(a)項に基づく、少なくても米国以外の1ケ国を指名したPCT国際出願の外国優先権を主張し、更に優先権の主張に係わる基礎出願の出願日前の出願日を有する外国特許出願、又は発明者証出願或るいはPCT国際出願を以下に"なし"の箱に印をつけることにより明記する：

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

METHOD FOR PREDICTING FUNCTIONS OF PROTEIN

_____

_____

the specification of which is attached hereto unless the following box is checked:

☒ was filed on May 26, 1998 as

United States Application Number _____

and was amended on _____ (if applicable) or,

PCT International Application Number PCT/JP98/02302

and was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority under Title 35, United States Code §119(a-d) or §365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT international application which designated at least one country other than the United States, listed below. I have also identified below, by checking the "No" box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed:

Prior foreign applications 先の外国出願 (Number) (番号)	(Country) (国名)	(Day/Month/Year Filed) (出願の年月日)	Priority claimed 優先権の主張 Yes あり	No なし
			☐	☐
			☐	☐

☐ その他の外国特許出願番号は別紙の追補優先権欄にて記載する。

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

Letter Version

Japanese Language Utility or Design Patent Application Declaration

私は、合衆国法典第35部第119条(e)項に基づく、下記の合衆国仮特許出願の利益を主張する。

I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

(Application No.) (出願番号)	(Day/Month/Year Filed) 出願の年月日
(Application No.) (出願番号)	(Day/Month/Year Filed) 出願の年月日
(Application No.) (出願番号)	(Day/Month/Year Filed) 出願の年月日

☐ その他の合衆国仮特許出願番号は別紙の追補優先権欄にて記載する。

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

私は、合衆国法典第35部第120条に基づく下記の合衆国特許出願、又は第365条(c)項に基づく合衆国を指名したPCT国際出願の利益を主張し、本願の請求の範囲各項に記載の主題が合衆国法典第35部第112条第1項規定の態様で、先の合衆国特許出願又はPCT国際出願に開示されていない限度において、先の出願の出願日と本願の国内出願日又はPCT国際出願日の間に有効となった連邦規則法典第37部第1章第56条に記載の特許要件に所要の情報を開示すべき義務を有することを認める。

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s), or §365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

(Application No.) (出願番号)	(Day/Month/Year Filed) (出願の年月日)	(現況) (特許済み、係属中　放棄済み)	(Status) (patented, pending, abandoned)
(Application No.) (出願番号)	(Day/Month/Year Filed) (出願の年月日)	(現況) (特許済み、係属中　放棄済み)	(Status) (patented, pending, abandoned)

☐ その他の合衆国又は国際特許出願番号は別紙の追補優先権欄にて記載する。

☐ Additional U.S. or international application numbers are listed on a supplemental priority sheet attached hereto.

私は、ここに自己の知識にもとずいて行った陳述がすべて真実であり、自己の有する情報および信ずるところに従って行った陳述が真実であると信じ、さらに故意に虚偽の陳述等を行った場合、合衆国法典第18部第1001条により、罰金もしくは禁錮に処せられるか、またはこれらの刑が併科され、またかかる故意による虚偽による陳述が本願ないし本願に対して付与される特許の有効性を損なうことがあることを認識して、以上の陳述を行ったことを宣言する。

私、下記署名者は、ここに記載の米国弁護士または代理人に本出願に関し特許商標庁にて取られるいかなる行為に関して、同米国弁護士又は代理人が、私に直接連絡なしに私の外国弁護士或るいは法人代表者からの指示を受け取り、それに従うようここに委任する。この指示を出す者が変更の場合には、ここに記載の米国弁護士又は代理人にその旨通知される。

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

The undersigned hereby authorizes the U.S. attorney or agent named herein to accept and follow instructions from either his foreign patent agent or corporate representative, if any, as to any action to be taken in the Patent and Trademark Office regarding this application without direct communication between the U.S. attorney or agent and the undersigned. In the event of a change in the persons from whom instructions may be taken, the U.S. attorney or agent named herein will be so notified by the undersigned.

Case # P20294

Japanese Language Utility or Design Patent Application Declaration

委任状： 私は、下記発明者として、下記に明記された顧客番号を伴う以下の弁護士又は、代理人をここに選任し、本順の手続きを遂行すること並びにこれに関する一切の行為を特許商標庁に対して行うことを委任する。そして全ての通信はこの顧客番号宛に発送される。

顧客番号 7055

現在選任された弁護士は下記の通りである。

Neil F. Greenblum	Reg. No. 28,394
Bruce H. Bernstein	Reg. No. 29,027
James L. Rowland	Reg. No. 32,674
Arnold Turk	Reg. No. 33,094

POWER OF ATTORNEY: As a named inventor, I hereby appoint the attorney(s) and/or agent(s) associated with the Customer Number provided below to prosecute this application and transact all business in the Patent and Trademark Office connected therewith, and direct that all correspondence be addressed to that Customer Number:

CUSTOMER NUMBER 7055

The appointed attorneys presently include:

Stephen M. Roylance	Reg. No. 31,296
William E. Lyddane	Reg. No. 41,568
William Pieprz	Reg. No. 33,630
Leslie J. Paperner	Reg. No. 33,329

Address: GREENBLUM & BERNSTEIN, P.L.C.
1941 ROLAND CLARKE PLACE
RESTON, VA 20191

直接電話連絡先：（名称および電話番号）

Direct Telephone Calls to: (*name and telephone number*)

GREENBLUM & BERNSTEIN, P.L.C.
(703) 716-1191

唯一のまたは第一の発明者の氏名 1-00	Full name of sole or first inventor Akiko ITAI
同発明者の署名 日付	Inventor's signature Akiko Itai Date February 16, 2001
住所	Residence Tokyo, Japan
国籍	Citizenship Japan
郵便の宛先	Post Office Address 5-16-6, Hongo, Bunkyo-ku, Tokyo 113-0033 Japan
第２の共同発明者の氏名（該当する場合） 2-00	Full name of second joint inventor, if any Nobuo TOMIOKA
同第２共同発明者の署名 日付	Second inventor's signature Date Nobuo Tomioka February 16, 2001
住所	Residence Tokyo, Japan JPX
国籍	Citizenship Japan
郵便の宛先	Post Office Address Room 302, Mansion Yuki, 5-21-12, Hongo,
	Bunkyo-ku, Tokyo 113-0033 Japan

（第六またはそれ以降の共同発明者に対しても同様な情報および署名を提供すること。）

(Supply similar information and signature for third and subsequent joint inventors.)

Case # P20294

Japanese Language Utility or DesignPatent Application Declaration

第三の共同発明者の氏名 3-00	Full name of third inventor, if any Reiko ITAI
共同発明者の署名 日付	Third inventor's signature Reiko Itai Date February 16, 2001
住所	Residence Tokyo, Japan JPX
国籍	Citizenship Japan
郵便の宛先	Post Office Address 5-16-6, Hongo, Bunkyo-ku, Tokyo 113 -0033 Japan
第四の共同発明者の氏名 4-00	Full name of fourth inventor, if any Masazumi IMAMURA
共同発明者の署名 日付	Fourth inventor's signature Masazumi Imamura Date February 16, 2001
住所	Residence Chiba, Japan JPX
国籍	Citizenship Japan
郵便の宛先	Post Office Address Room 301, Garden-Court Morinomachi-5-bankan, 39, Asumigaoka 4-chome, Midori-ku, Chiba-shi, Chiba 267-0066 Japan
第五の共同発明者の氏名	Full name of fifth inventor, if any
共同発明者の署名 日付	Fifth inventor's signature Date
住所	Residence
国籍	Citizenship
郵便の宛先	Post Office Address
第六の共同発明者の氏名	Full name of sixth inventor, if any
共同発明者の署名 日付	Sixth inventor's signature Date
住所	Residence
国籍	Citizenship
郵便の宛先	Post Office Address

(それ以降の共同発明者にたいしても同様な情報および署名を提供すること。)

(Supply similar information and signature for subsequent joint inventors.)